

ANALISIS DISKRIMINAN FISHER POPULASI GANDA UNTUK KLASIFIKASI NASABAH KREDIT

Ungu Siwi Maharunti¹, Moch. Abdul Mukid², Agus Rusgiyono³

¹Mahasiswa Jurusan Statistika FSM Universitas Diponegoro

^{2,3}Staff Pengajar Jurusan Statistika FSM Universitas Diponegoro

e-mail ungusiwi@gmail.com

ABSTRACT

Credit is the biggest asset carried out by a bank and become the most dominant contributor to the bank income. However, the activity to distribute the credit takes a risk which can influence health and continuance of bank business. The credit risk which potentially occurs can be measured and controlled by analyzing directly whichever the credit client categorized to. The credit risk categorized to current credit, in specific concern credit, less current credit, doubtful credit and bad credit based on Bank Indonesia Regulation No.: 7/2/PBI/2005. The independent variables used in this research are nominal credit, principal balance, in time being bank client, time period, and bank interest. Fisher multiple discriminant analysis is a method whose assumption equality of covariance matrices. The result from using the Fisher multiple discriminant analysis in data of credit client from bank "X" in Pati shows that variable principal balance, in time being bank client, time period, and bank interest significant to measure credit risk. The classification using the Fisher multiple discriminant analysis in data of credit client from bank "X" in Pati gives the accurate 64,33%.

Keywords: credit, classification, fisher multiple discriminant analysis

1. PENDAHULUAN

Dalam UU Nomor 10 Tahun 1998 tentang Perubahan atas UU Nomor 7 Tahun 1992 tentang Perbankan, disebutkan bahwa kredit adalah penyediaan uang atau tagihan, berdasarkan persetujuan atau kesepakatan pinjam meminjam antara bank dengan pihak lain yang mewajibkan pihak peminjam untuk melunasi utangnya setelah jangka waktu tertentu dengan pemberian bunga. Kegiatan menyalurkan kredit mengandung risiko yang dapat mempengaruhi kesehatan dan kelangsungan usaha bank. Dalam pelaksanaannya terdapat sejumlah permasalahan yang muncul dari program kredit pinjaman. Salah satu permasalahan yang sering muncul adalah adanya nasabah yang telat membayar angsuran. Oleh karena itu, setiap bank harus menjaga kualitas kreditnya sebaik mungkin dan sedini mungkin untuk mengenali kemunculan penurunan kualitas kredit.

Menurut Johnson dan Wichern (2007), analisis diskriminan digunakan untuk mengklasifikasikan individu ke dalam salah satu dari dua kelompok atau lebih. Fungsi diskriminan yang dihasilkan dapat digunakan untuk memberikan pendugaan yang paling tepat untuk mengklasifikasikan individu ke dalam salah satu kelompok berdasarkan skor variabel bebas (skor diskriminan). Misalnya untuk mengklasifikasikan debitur ke dalam kelompok debitur yang tidak menunggak atau kelompok debitur yang menunggak. Pada data keuangan, khususnya data kredit, variabel-variabel bebas yang digunakan sering tidak mengikuti distribusi normal. Oleh karena itu, pada tulisan ini akan dilakukan pengklasifikasian nasabah kredit menggunakan analisis diskriminan Fisher populasi ganda. Analisis diskriminan Fisher populasi ganda adalah metode yang dikembangkan oleh R. A Fisher (1936) yang memerlukan asumsi kesamaan matriks varian-kovarian namun tidak mengharuskan variabel bebas berdistribusi normal multivariat. Tujuan yang ingin dicapai dalam penelitian ini adalah mengidentifikasi variabel-variabel yang berpengaruh terhadap

status kredit nasabah menggunakan klasifikasi diskriminan Fisher populasi ganda dan mengetahui akurasi ketepatan klasifikasi menggunakan analisis diskriminan Fisher populasi ganda untuk klasifikasi nasabah kredit.

2. TINJAUAN PUSTAKA

2.1. Pengertian Bank

Pengertian bank menurut yang telah dituangkan dalam Undang-undang Negara Republik Indonesia pada nomor 10/1998 pasal 1 huruf 2 di mana telah mengatur mengenai perbankan yang menjelaskan bahwa pengertian bank adalah badan usaha yang menghimpun dana dari masyarakat dalam bentuk simpanan dan menyalurkannya kepada masyarakat dalam bentuk kredit atau bentuk-bentuk lainnya dalam rangka meningkatkan taraf hidup rakyat banyak.

2.2. Pengertian Nasabah

Definisi nasabah menurut Undang-Undang No. 10 Tahun 1998 tentang Perubahan Atas Undang-Undang No. 7 Tahun 1992 tentang Perbankan diatur perihal nasabah yang terdiri dari dua pengertian yaitu:

1. Nasabah penyimpan adalah nasabah yang menempatkan dananya di bank dalam bentuk simpanan berdasarkan perjanjian bank dengan nasabah yang bersangkutan.
2. Nasabah debitur adalah nasabah yang memperoleh fasilitas kredit atau pembiayaan berdasarkan prinsip syariah atau yang dipersamakan dengan itu berdasarkan perjanjian bank dengan nasabah yang bersangkutan.

2.3. Pengertian Kredit

Menurut UU. No. 10 Tahun 1998, pengertian kredit adalah suatu penyediaan uang atau tagihan yang dapat dipersamakan dengan itu, berdasarkan persetujuan atau kesepakatan pinjam-meminjam antara bank dengan pihak lain yang mewajibkan pihak peminjam untuk melunasi hutangnya setelah jangka waktu tertentu dengan pemberian bunga. Istilah kredit berasal dari bahasa latin yaitu *credere* yang berarti kepercayaan, atau *credo* yang berarti saya percaya, artinya kepercayaan dari kreditur (pemberian pinjaman) bahwa debiturnya (penerima pinjaman) akan mengembalikan pinjaman beserta bunganya sesuai dari perjanjian kedua belah pihak. Sedangkan dalam pasal 12 Ayat 3, kualitas kredit ditetapkan kedalam lima kategori yaitu: a. Lancar; b. Dalam Perhatian Khusus; c. Kurang Lancar; d. Diragukan; atau e. Macet.

2.4. Analisis Diskriminan

Analisis diskriminan adalah teknik multivariat untuk memisahkan objek-objek dalam kelompok yang berbeda dan mengelompokkan objek baru ke dalam kelompok-kelompok tersebut (Johnson, 2007). Tujuan utama dari *multiple discriminant analysis* adalah untuk mengetahui perbedaan antar kelompok. Umumnya peubah respon terdiri dari dua klasifikasi namun pada beberapa kasus peubah respon memiliki lebih dari dua klasifikasi. Jika terdapat hanya dua klasifikasi maka disebut analisis diskriminan dua kelompok namun jika terdapat lebih dari dua disebut *Multiple Discriminant Analysis* (MDA) (Hair et al., 1984).

Menurut Hair et al. (1984), analisis diskriminan melibatkan kombinasi dari dua atau lebih variabel independen yang akan membentuk pemisah terbaik di antara kelompok

awal. Pada analisis diskriminan setiap unit amatan akan memiliki p nilai karakteristik yang merupakan nilai dari p variabel prediktor yang digunakan.

2.5. Pemilihan Variabel Bebas

Terdapat dua pendekatan dalam pemilihan variabel bebas yang akan digunakan untuk menduga koefisien diskriminan, yaitu direct method dan stepwise discriminant analysis. Direct method meliputi estimasi koefisien fungsi diskriminan dimana seluruh variabel bebas terlibat, secara simultan dimasukkan bersama-sama dalam analisis tanpa memperlihatkan discriminating power. Metode yang kedua yaitu stepwise discriminant analysis, variabel bebas (prediktor) dimasukkan secara berurutan didasarkan pada kemampuannya untuk mendiskriminasi antar-kelompok (Supranto, 2010).

Prosedur stepwise dilakukan dengan memilih variabel yang memiliki kemampuan diskriminan terbaik. Jarak Mahalanobis (D2) digunakan dalam perhitungan metode stepwise untuk menentukan variabel yang memiliki kemampuan terbaik dalam mendiskriminasi kelompok. Jarak Mahalanobis didefinisikan dengan rumus

$$d_{AB} = \sqrt{(x_A - x_B)' S_{AB}^{-1} (x_A - x_B)}$$

dengan

x_A = nilai amatan pada kelompok ke-A

x_B = nilai amatan pada kelompok ke-B

S_{AB} = matriks varian kovarian

Variabel bebas saling dipasangkan kemudian variabel yang memiliki nilai signifikan dan memiliki jarak mahalanobis terbesar dipilih sebagai variabel yang memiliki kemampuan memisahkan populasi dan dianggap layak untuk dimasukkan ke dalam fungsi diskriminan (Hair et al., 1984).

2.6. Fungsi Diskriminan Linier Fisher untuk Populasi Ganda

Pendekatan Fisher tidak mengharuskan kenormalan multivariat pada variabel-variabel penjelasnya namun mengasumsikan kesamaan matriks varian kovarian (Johnson, 2007). Misalkan dalam suatu populasi yang terdiri dari g populasi dan sejumlah n_i observasi untuk setiap kelompok ke- i , dengan setiap observasi sebanyak p variabel bebas atau p atribut $\mathbf{X}' = (X_1, X_2, \dots, X_p)$. Fisher mengelompokkan suatu observasi berdasarkan nilai skor yang dihitung dari suatu kombinasi linier sebagai berikut:

$$Y = \mathbf{a}' \mathbf{X}$$

dengan

Y = fungsi diskriminan

\mathbf{a} = koefisien fungsi diskriminan

\mathbf{X} = variabel prediktor

Nilai harapan untuk Y , atau dituliskan $E(Y) = \mathbf{a}' E(\mathbf{X} | \pi_i) = \mathbf{a}' \mu_i$ untuk setiap populasi π_i dan varian $(Y) = \mathbf{a}' \text{Cov}(\mathbf{X}) \mathbf{a} = \mathbf{a}' \Sigma \mathbf{a}$ untuk semua populasi, dengan Σ merupakan matriks varian kovarian dari \mathbf{X} . Akibatnya, nilai harapan dari $\mu_{iY} = \mathbf{a}' \mu_i$ berubah sesuai dengan populasi dari \mathbf{X} yang dipilih. Dalam hal ini, fungsi diskriminan Fisher memaksimalkan suatu rasio

$$\frac{(\text{total kuadrat jarak dari populasi ke rata – rata keseluruhan dari } Y)}{(\text{varians dari } Y)}$$

$$= \frac{\sum_{i=1}^g (\mu_{iY} - \bar{\mu}_Y)^2}{\sigma_Y^2} = \frac{\sum_{i=1}^g (a' \mu_i - a' \bar{\mu})^2}{a' \Sigma a} = \frac{a' (\sum_{i=1}^g (\mu_i - \bar{\mu})(\mu_i - \bar{\mu})') a}{a' \Sigma a}$$

$$\text{atau bisa dikatakan } \frac{\sum_{i=1}^g (\mu_{iY} - \bar{\mu}_Y)^2}{\sigma_Y^2} = \frac{a' B_{\mu} a}{a' \Sigma a}.$$

Pendugaan dari Σ berdasarkan masing-masing sampel ditunjukkan dengan matriks $W = \sum_{i=1}^g (n_i - 1) S_i = \sum_{i=1}^g \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)(x_{ij} - \bar{x}_i)'$.

Dengan demikian $W/(n_1 + n_2 + \dots + n_g - g) = S_{pooled}$ merupakan penduga dari Σ . Supaya dapat memisahkan kelompok secara maksimal, fungsi diskriminan Y diduga dengan kriteria memaksimumkan variabilitas antar kelompok. Koefisien a yang membuat $\frac{a' B a}{a' S_{pooled} a}$ maksimum, juga memaksimumkan $\frac{a' B a}{a' W a}$. Selanjutnya, diperoleh nilai a maksimum dengan bentuk vektor eigen \hat{e}_i dari $W^{-1} B$.

Aturan klasifikasi diskriminan Fisher adalah menempatkan y ke dalam populasi π_k jika kuadrat jarak dari y ke μ_{kY} lebih kecil dibandingkan kuadrat jarak dari y ke μ_{iY} , untuk $i \neq k$.

Misalkan digunakan fungsi diskriminan sebanyak r maka alokasikan x ke dalam populasi π_k jika

$$\sum_{j=1}^s (y_j - \bar{y}_{kj})^2 \leq \sum_{j=1}^r [a_j' (x - \bar{x}_i)]^2, i \neq k$$

$$\text{dengan } \sum_{j=1}^s (y_j - \bar{y}_{kj})^2 = \sum_{j=1}^r [a_j' (x - \bar{x}_k)]^2$$

j = banyaknya fungsi yang terbentuk ($j=1,2,\dots,s$)

k = banyaknya kelompok

Dengan demikian, alokasikan x ke dalam populasi π_k dimana k menghasilkan nilai minimum dari $\sum_{j=1}^s (y_j - \bar{y}_{kj})^2$.

2.7. Ketepatan Fungsi Klasifikasi

Menurut Johnson (2007), *apparent error rate (APER)* menyatakan bahwa secara ekuivalen kinerja sebuah model dapat dinyatakan dalam bentuk *error rate*-nya.

Nilai APER ditunjukkan dengan

$$APER = \frac{N - \sum_{i=1}^g n_{ii}}{N}$$

dengan

N = banyaknya pengamatan pada semua kelompok

n_{ii} = banyaknya pengamatan yang tepat diklasifikasikan dari kelompok aktual ke- i pada kelompok prediksi ke- i

g = banyaknya kelompok

3. METODOLOGI PENELITIAN

3.1. Data

Data yang digunakan dalam penelitian ini merupakan data sekunder yang diperoleh dari data debitur di suatu bank pada tahun 2015. Unit observasi pada penelitian ini adalah kota Pati Provinsi Jawa Tengah.

Variabel respon dalam penelitian ini merupakan status kredit dari nasabah kredit di suatu bank dengan kategori 1 untuk status kredit lancar, kategori 2 untuk status dalam

perhatian khusus, kategori 3 untuk status kurang lancar, kategori 4 untuk status diragukan, dan kategori 5 untuk status macet. Variabel prediktor dalam penelitian ini ada 5, yaitu :

1. Nominal kredit yang diajukan nasabah (X1)
2. Saldo rekening yang dimiliki oleh nasabah (X2)
3. Suku bunga yang dibebankan kepada nasabah (X3)
4. Jangka waktu pembayaran kredit dalam satuan bulan (X4)
5. Lama menjadi nasabah bank dalam satuan bulan (X5)

3.2. Metode Analisis

Metode analisis yang digunakan dalam penelitian ini menggunakan analisis diskriminan Fisher dengan langkah-langkah sebagai berikut :

1. Uji homogenitas varian kovarian
2. Membagi data menjadi data latih dan data uji
3. Menduga fungsi diskriminan dengan menggunakan fungsi diskriminan linier Fisher
4. Melakukan klasifikasi
5. Uji ketepatan model

4. HASIL DAN PEMBAHASAN

4.1. Analisis Diskriminan Fisher Populasi Ganda

4.1.1. Uji Kesamaan Matriks Varian Kovarian

Asumsi yang diperlukan ketika membandingkan dua atau lebih vektor mean adalah matriks kovarian dari populasi yang berbeda tidak berbeda nyata atau dapat dikatakan sama. Pengujian yang dilakukan adalah dengan uji Box'M. Hipotesisnya sebagai berikut.

$$H_0 : \Sigma_1 = \Sigma_2 = \Sigma_3 = \Sigma_4 = \Sigma_5 = \Sigma \text{ (Matriks kovarian tidak berbeda nyata)}$$

$$H_1 : \Sigma_i \neq \Sigma_j \text{ dengan } i,j = 1,2,\dots,5 \text{ (Setidaknya ada dua matriks kovarian yang berbeda)}$$

Taraf Signifikansi : $\alpha = 5\%$

Statistik Uji :

$$\text{Nilai Box's M} = 56,236$$

$$p\text{-value} = 0,138$$

Kriteria Uji :

H_0 ditolak apabila nilai $p\text{-value} < \alpha$

Keputusan :

H_0 gagal ditolak karena nilai signifikansi = $0,138 > 0,05$

Kesimpulan :

Sehingga dapat disimpulkan bahwa matriks varian kovarian untuk semua populasi tidak berbeda nyata.

4.1.2. Pemilihan Variabel Bebas

Pada penelitian ini, digunakan metode *stepwise discriminant* dalam memilih variabel bebas yang masuk ke dalam model diskriminan. Prosedur *stepwise* dilakukan dengan memilih variabel yang memiliki kemampuan diskriminan terbaik. Jarak Mahalanobis (D^2) antar dua kelompok yang dihasilkan oleh variabel bebas digunakan dalam perhitungan metode *stepwise* untuk menentukan variabel yang memiliki kemampuan terbaik dalam mendiskriminasi kelompok. Namun sebelum melihat nilai D^2 , variabel bebas tersebut harus memenuhi kriteria signifikan dengan nilai signifikansinya kurang dari 0,05. Setelah dilakukan tahapan *stepwise method* diperoleh hasil bahwa variabel yang signifikan

terhadap status kredit nasabah dan dapat digunakan ke dalam fungsi diskriminan adalah X_2 , X_3 , X_4 , dan X_5 .

4.1.3. Pendugaan Fungsi Diskriminan

Berdasarkan perhitungan dengan variabel yang signifikan, diketahui matriks jumlah kuadrat dan jumlah *cross product* antar kelompok sampel dengan

$$B = \sum_{i=1}^5 (\bar{x}_i - \bar{x})(\bar{x}_i - \bar{x})'$$

$$= \begin{bmatrix} 11360,74332 & 93,319364 & -356,3719 & -189,097 \\ 93,31936 & 11,106222 & -31,938672 & -1,601614 \\ -356,37193 & -36,300613 & 53,848438 & 1,626434 \\ -189,09744 & -1,601614 & 1,626434 & 33,788689 \end{bmatrix}$$

$$W = \begin{bmatrix} 1981464,387 & 5375,97404 & 3621,35 & 8284,56683 \\ 5375,974 & 1612,84391 & 132,0346 & 76,94215 \\ 3621,350 & 132,03462 & 3020,8667 & 299,07756 \\ 8284,567 & 76,94215 & 299,0776 & 2964,33149 \end{bmatrix}$$

$$W^{-1}B = \begin{bmatrix} 0,006109035 & 0,000047276 & -0,000179368 & -0,0001414 \\ 0,050708523 & 0,007729393 & -0,0253413406 & -0,001053 \\ -0,120578063 & -0,01244904 & 0,0434845112 & -0,0004206 \\ -0,070014968 & 0,0003829658 & -0,002679526 & 0,01186338 \end{bmatrix}$$

Fungsi diskriminan yang terbentuk sebagai berikut :

$$Y_1 = 0,000X_2 - 0,110X_3 + 0,194X_4 - 0,110X_5$$

$$Y_2 = -0,003X_2 - 0,056X_3 + 0,083X_4 + 0,195X_5$$

$$Y_3 = 0,008X_2 - 0,066X_3 + 0,005X_4 + 0,046X_5$$

$$Y_4 = 0,001X_2 + 0,275X_3 + 0,082X_4 + 0,002X_5$$

4.2. Ketepatan Hasil Klasifikasi

4.2.1 Ketepatan Hasil Klasifikasi Data Latih

$$APER = \frac{N - \sum_{i=1}^5 n_{ii}}{N} = \frac{157 - 101}{157} = 0,3567$$

Setelah dilakukan perhitungan, diperoleh nilai APER sebesar 0,3567 atau 35,67% dan dengan demikian ketepatan klasifikasinya sebesar 0,6433 atau 64,33% data tepat diklasifikasikan.

4.2.2 Ketepatan Hasil Klasifikasi Data Uji

$$APER = \frac{N - \sum_{i=1}^5 n_{ii}}{N} = \frac{67 - 46}{67} = 0,3125$$

Setelah dilakukan perhitungan, diperoleh nilai APER sebesar 0,3134 atau 31,34% dan dengan demikian ketepatan klasifikasinya sebesar 0,6866 atau 68,66% data tepat diklasifikasikan.

5. KESIMPULAN

Berdasarkan pembahasan yang telah dipaparkan pada bab IV, maka dapat diambil kesimpulan sebagai berikut :

1. Variabel yang berpengaruh terhadap status kredit nasabah bank “X” menggunakan klasifikasi model diskriminan Fisher populasi berganda adalah variabel saldo

rekening yang dimiliki nasabah, suku bunga yang dibebankan, jangka waktu pembayaran kredit dan lama menjadi nasabah bank tersebut.

2. Persentase ketepatan klasifikasi data latih nasabah kredit bank “X” di Kota Pati sebesar 64,33% data tepat diklasifikasikan dengan banyaknya data keanggotaan aktual yang tepat diklasifikasikan ke dalam keanggotaan prediksi sebanyak 101 data dan banyaknya data keanggotaan aktual yang salah diklasifikasikan ke dalam keanggotaan prediksi sebanyak 56 data.
3. Persentase ketepatan klasifikasi data latih nasabah kredit bank “X” di Kota Pati sebesar 68,66% data tepat diklasifikasikan dengan banyaknya data keanggotaan aktual yang tepat diklasifikasikan ke dalam keanggotaan prediksi sebanyak 46 data dan banyaknya data keanggotaan aktual yang salah diklasifikasikan ke dalam keanggotaan prediksi sebanyak 21 data.

DAFTAR PUSTAKA

- Dillon, W. and Goldstein, M. 1984. *Multivariate Analysis Methods and Application*. New York: John Wiley&Sons, Inc.
- Hair, J.F., Anderson R.E., Tatham R.L., and Black W.C. 1998. *Multivariate Data Analysis. Fifth Edition*. Pearson Education Prentice Hall. Inc.
- Johnson, R.A. and Wichern. D.W. 2007. *Applied Multivariate Statistical Analysis. Sixth Edition*. New Jersey: Prentice Hall International. Inc.
- Leon, B., Ericson S. 2007. *Manajemen Aktiva Pasiva Bank Nondevisa*. Jakarta: PT Grasindo.
- Supranto.2010. *Analisis Multivariat Arti & Interpretasi*. Jakarta: PT Rineka Cipta.
- Suyatno. T. 1988. *Kelembagaan Perbankan*. Jakarta: PT Gramedia Pustaka Utama.
- Rencher, A.C. 2002. *Methods of Multivariate Analysis. Second Edition*. Canada: John Wiley & Sons Inc.